Anamaria Crisan, PhD

amcrisan amcrisan amcrisan
https://amcrisan.github.io

I am a Lead Research Scientist at Tableau (a Salesforce company), where I am part of the Office of the Chief Technology Officer (OCTO). I have Ph.D. in Computer Science and 7 years of experience translating applied machine learning (ML) research into products and policy-making within industry, government, and academia. My research track record includes award-winning publications at top-tier conferences of the ACM (CHI, FAccT) and IEEE (TVCG, CG&A) in addition to biomedical journals (Nature, Oxford Bioinformatics, PLOS). My research goals are to help people make informed and trustworthy decisions with data.

Interests: Interactive Machine Learning, Human Centered AI/ML, Data Science, Data Visualization, Bioinformatics

Education

2015 - 2019 **PhD Computer Science** University of British Columbia
Thesis Supervisors: Dr. Tamara Munzner and Dr. Jennifer L. Gardy
Notable Awards: Canada Vanier Scholar and UBC Public Scholar

2008 - 2010 **MSc Bioinformatics** University of British Columbia
Thesis Supervisors: Dr. Sohrab Shah and Dr. Samuel Aparicio
Notable Awards: CHIR-MSFHR Bioinformatics Trainee Award

2004 - 2008 **BSc Computer Science** Queen's University at Kingston
Honors Thesis Supervisor: Dr. Hagit Shatkay
Notable Awards: NSERC USRA, CRA-W Distributed Mentorship Trainee Award

Research Positions

My research addresses the ways people and ML/AI trade information and decide upon boundaries in collaborative decision-making. My methods are rooted in my experiences as a biomedical researcher working alongside multidisciplinary teams and wrestling with the challenges of applying AI/ML to high-risk decision-making.

2019 - Present **Lead Research Scientist, Tableau Research**
Seattle, Washington, USA

- I develop strategic research directions for human-ML/AI collaboration, including technical approaches and best practice toward responsible use of ML/AI
- I lead research projects in partnership with cross-functional engineering and product teams to inform, influence, and support ML/AI product strategy
- I collaborate closely with UR/UX to create human-centered ML/AI experiences
- I create and analyze complex datasets to solve challenging problems using different statistical and ML approaches
- I contribute to the Tableau Main (`Overcoming Misinformation`, `Data Science Reflection`) and Engineering Blog Sites (`Interactive Machine Learning`, `AutoML and People`)

2013 - 2015

Research Data Scientist (Bioinformatics), British Columbia Centre for Disease Control
Vancouver, British Columbia, CANADA

- Routinely collaborated with doctors, nurses, health officers, and statisticians to study, model, and monitor spread disease using genomic technology + ML
- Developed data visualization tools to help domain experts get started with genomic + electronic health record data more effectively. Developed the reportable disease dashboard prototype that now appears on the BCCDC website.
- Developed and deployed ML models for improved water quality measures from metagenomic signatures from wastewater sampling
- Developed a clinical report for TB that has been adopted by global health agencies
- Routinely invited to give talks about my work, and its findings, at medical grand rounds, to Canadian government officials, and also global public health agencies
- Conducted statistical analysis of data, including linear (and hierarchical) modelling, time series analysis, spatio-temporal analysis, causal analysis, clustering, supervised and unsupervised learning

2010 - 2013

Research Data Scientist (Bioinformatics), Decipher Biosciences
Vancouver, British Columbia, CANADA

- Start-up (then GenomeDx Biosciences) where I was employee 3, grew it up to 100 people, and which sold recently with a 600 million valuation.
- I was the co-inventor of the flagship product : a genomic classifier that predicts metastatic prostate cancer post-surgery. This product is used in clinics today.
- I participated in a series A funding round valued at around 12 million; I worked directly with investors on the technical due diligence
- I directly to contributed research and clinical study development for our portfolio for medicare reimbursement
- Collaborated with clinical and research partners to conduct statistical analysis of data, design, and analysis of clinical studies

Selected Publications

I have a total of 27 publications, with over 3,700 citations.
A full publication list with updated citation counts are available on my Google Scholar profile

FAccT 2022

Crisan A., Drouhard M., Vig J, Ranjani, N., *Interactive Model Cards: Human Centered Approach to Model Documentation.* Proc. FAccT'22
10.1145/3531146.3533108 Python Streamlit App

TVCG 2021

Crisan A, Fisher, S†, Gardy, JL, Munzner T., *GEViTRec: Data reconnaissance through recommendation using a domain-specific visualization prevalence design space.* IEEE Transaction on Visualization and Computer Graphics
10.1109/TVCG.2021.3107749

CHI 2021

Crisan A, Fiore-Gartland, B., *Fits and Starts: Enterprise use of AutoML and the role of humans in the loop.* Proc. of CHI'21
10.1145/3411764.3445775 CHI'21 Video Tablaeu Engineering Blog
Best Paper Honorable Mention - CHI'21

CHI 2021

Crisan A, Correll M.,*User ex Machina: Simulation as a design probe in human in the loop text analytics.* Proc. of CHI'21
10.1145/3411764.3445425 CHI'21 Video Tableau Engineering Blog

TVCG 2021 **Crisan A.**, Fiore-Gartland, B., Tory, M. *Passing the data baton: A retrospective analysis on Data Science work and workers.* IEEE Transactions on Visualization and Computer Graphics
10.1109/TVCG.2020.3030340 VDS'21 Video Tableau Main Blog
Best Paper at the Visualization in Data Science Symposium

Bioinf. 2018 **Crisan A.**, Munzner T., Gardy JL., *Adjutant: an R-based tool to support topic discovery for systematic and literature reviews* Oxford Bioinformatics
10.1093/bioinformatics/bty722 R Package

PLOS 2013 **Crisan A.**, Erho N., Vergara IA., Mitra AP., Ghadessi M., Buerki C., Bergstralh EJ., Kollmeyer T., Fink S., Haddad Z., Zimmermann B., Sierocinski T., Ballman KV., Triche TJ., Black PC., Karnes RJ., Klee G., Davicioni E., Jenkins RB., *Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy.* PLOS ONE
10.1371/journal.pone.0066855

Selected Presentations

CHIL 2022 Are Log Scales Endemic Yet? Strategies for Visualizing Biomedical and Public Health Data
Location: Virtual (due to COVID-19)

VIZBI 2022 Visualization in Public Health
Location: Virtual (due to COVID-19)

KDD 2021 Natto: Rapid Visual Iteration of Analytic Data Models with Intelligent Assistance
Location: Virtual (due to COVID-19)

CHI 2021 Fits and Starts: Enterprise Use of AutoML and the Role of Humans in the Loop
Vis ex Machina: Simulation as a Design Probe in Human in the Loop Text Analytics
Location: Virtual (due to COVID-19)

ABPHM 2019 Automated Visualization Recommendations for Genomic Epidemiology
Location: Hinxton, UK
Winner - ABPHM'19 Software Prize

Dagstuhl 2019 Dagstuhl Seminar 18161 – BioVis Crossroads
Location: Schloss Dagstuhl, Germany
Organized by: Jan Aerts (KU Leuven, BEL); Nils Gehlenborg (Harvard University, USA);Elisabeta Marai (University of Illinois,USA); and Kay Nieselt (Uni. Tübingen, DEU)

Selected Awards

2021 Best Paper Honorable Mention - CHI'21
2020 Best Paper VDS@IEEE VIS'21
2015 Canada Vanier Scholar ($150,00)
2015 UBC Public Scholar ($5,00)
2017 UBC Affiliated Fellowship ($25,00)

Selected Service

Program Committees

2021- IEEE VIS PC
2020- Frontiers in Bioinformatics Review Editor
2017 - 2019 Bioinformatics Open Source Conference

Organizing Committees

2021-	IEEE VIS - VIS in Practice co-chair
2021-	VDS Papers co-chair
2018 - 2020	Biovis Challenges
2016 - 2018	Biovis website chair

Summary of Skills & Technical Competencies

- Experience in industrial research (5 years), government (2 years), and academia
- Track record of award winning high-quality research publications accepted to top-tier venues
- Demonstrated ability to plan, design, execute, actualize, and communicate both strategic and tactical translational ML/AI research to technical and non-technical audiences
- Proficiency in quantitative, qualitative, and mixed research methods. Expertise with advanced statistical data analysis, clinical study design, user research study design
- Proficiency with ML/AI and Data Science tooling in Python (pandas, sklearn, nltk, hugging face, fastai, etc.), R (ggplot, tidyverse, shiny,etc.), SQL, AWS (Sagemaker)
- Proficiency with Data Visualization libraries in JavaScript (Vega-lite, D3), Python (altair, streamlit) and R (ggplot, shiny), also software systems (Tableau)